Capacity of the State-Dependent Half-Duplex Relay Channel Without Source-Destination Link

Nikola Zlatanov and Robert Schober

Abstract

We derive the capacity of the state-dependent half-duplex relay channel without source-destination link. The output of the state-dependent half-duplex relay channel depends on the randomly varying channel states of the source-relay and relay-destination links, which are known causally at all three nodes. For this channel, we prove a converse and show the achievability of the capacity based on a buffer-aided relaying protocol with adaptive link selection. This protocol chooses in each times slot one codeword to be transmitted over either the source-relay or the relay-destination channel depending on the channel states. Our proof of the converse reveals that state-dependent half-duplex relay networks offer one additional degree of freedom which has been previously overlooked. Namely, the freedom of the half-duplex relay to choose when to receive and when to transmit.

I. Introduction

The capacities of the memoryless full-duplex and half-duplex relay channels without source-destination link were derived in [1] and [2], respectively. For the case when the links undergo fading and the nodes have full channel state information (CSI), only the capacity of the fading full-duplex relay channel without source-destination link is known [3]. In contrast, for the fading half-duplex relay channel without source-destination link only achievable rates are known as presented in [3] and [4], and recently in [5]. However, for the achievable rates presented in [3] and [4], the relay is assumed to always alternate between receiving in one time slot and transmitting in the following time slot, which for fading may lead to a large rate loss as shown in

N. Zlatanov is with the Department of Electrical and Computer Engineering, University of British Columbia (UBC), Vancouver, BC, V6T 1Z4, Canada, email: zlatanov@ece.ubc.ca

R. Schober is the Chair for Digital Communication at the Friedrich-Alexander University (FAU), Erlangen, Germany, email: schober@lnt.de

[5]. In [5], a new achievable rate was introduced for the fading half-duplex relay channel without source-destination link. This new achievable rate was obtained with buffer-aided relaying and adaptive link selection. In this protocol, the relay has the freedom to choose when to receive from the source and when to transmit to the destination based on the fading states of the source-relay and relay-destination links. In [6], it was shown that adaptive link selection is also beneficial if the source and/or the relay transmit with fixed rates. The rate in [5] was derived only for the additive white Gaussian noise (AWGN) channel with time-continuous fading. Moreover, only an achievable rate was given and it is not clear from [5] whether or not higher rates are possible. In this paper, we consider the state-dependent half-duplex relay channel which is more general then the channel considered in [5]. For the state-dependent half-duplex relay channel without source-destination link, we prove a converse and show the achievability of the capacity for the case when the channel states change slowly. Thereby, we show that the rate achieved in [5] is the capacity rate. Furthermore, the converse proof reveals that in state-dependent half-duplex channels there is one additional degree of freedom, which has been previously overlooked, e.g. [3], [4]. This additional degree of freedom is due to the ability of a half-duplex node to choose when to receive and when to transmit. To this end, the relay has to have a buffer for information storage.

The rest of this paper is organized as follows. In Section II, we present the channel model. In Section III, we introduce the capacity of the channel and investigate three specific cases. In Section IV, we prove the converse and show the achievability of the derived capacity. Finally, Section IV concludes the paper.

II. System Model, Notations, and Definitions

The state-dependent half-duplex relay channel without source-destination link is shown in Fig. 1. It consists of a source, a half-duplex relay, and a destination. The half-duplex relay cannot transmit and receive at the same time. There is no link between the source and the destination, and therefore the source transmits its information to the destination through the relay. The source-relay and relay-destination channels are assumed to undergo different states, respectively, which are modeled as ergodic and stationary random processes.

The discrete memoryless state-dependent half-duplex relay channel is defined by $\mathcal{X}_1^*$, $\mathcal{X}_2^*$, $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{Y}_1^*$, $\mathcal{Y}_2^*$, $p(s_1, s_2)$, and $p(y_1, y_2|x_1, x_2, s_1, s_2)$, where $\mathcal{X}_1^*$ and $\mathcal{X}_2^*$ are the finite input alphabets

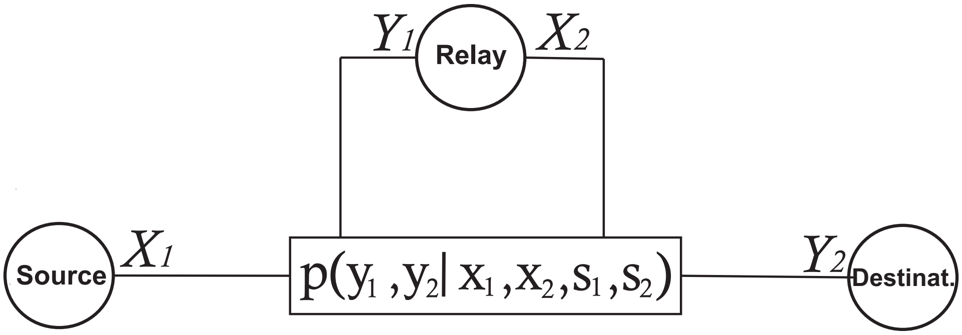

Fig. 1. System model comprising a source, a half-duplex relay equipped with a buffer, and a destination.

at the encoders of the source and the relay, respectively, $\mathcal{S}_1$ and $\mathcal{S}_2$ are the finite alphabets of the states governing the source-relay and relay-destination channels, respectively, $\mathcal{Y}_1^*$ and $\mathcal{Y}_2^*$ are the finite output alphabets at the decoders of the relay and destination, respectively, $p(s_1, s_2)$ is the probability mass function (PMF) on $\mathcal{S}_1 \times \mathcal{S}_2$, and $p(y_1, y_2|x_1, x_2, s_1, s_2)$ is the PMF on $\mathcal{Y}_1^* \times \mathcal{Y}_2^*$, for given $x_1 \in \mathcal{X}_1^*$, $x_2 \in \mathcal{X}_2^*$, $s_1 \in \mathcal{S}_1$, and $s_2 \in \mathcal{S}_2$. Thus, the probability of having states $s_1 \in \mathcal{S}_1$ and $s_2 \in \mathcal{S}_2$ on the source-relay and relay-destination channel, respectively, is $p(s_1, s_2)$, whereas the probability of the relay and destination received symbols $y_1 \in \mathcal{Y}_1^*$ and $y_2 \in \mathcal{Y}_2^*$, respectively, given that the source-relay and relay-destination channels are in states s_1 and s_2, respectively, and the source and relay transmitted symbols are $x_1 \in \mathcal{X}_1^*$ and $x_2 \in \mathcal{X}_2^*$, respectively, is $p(y_1, y_2|x_1, x_2, s_1, s_2)$. Since we assume a half-duplex relay, which cannot receive and transmit at the same time, the source's and relay's input alphabets are given by $\mathcal{X}_1^* = \mathcal{X}_1 \cup \{\varnothing\}$ and $\mathcal{X}_2^* = \mathcal{X}_2 \cup \{\varnothing\}$, respectively, where $\varnothing$ is a special symbol, distinct from those in $\mathcal{X}_1$ and $\mathcal{X}_2$, which denotes no input and is used because of the half-duplex constraint. On the other hand, the relay's and destination's output alphabets are given by $\mathcal{Y}_1^* = \mathcal{Y}_1 \cup \{\varnothing\}$ and $\mathcal{Y}_2^* = \mathcal{Y}_2 \cup \{\varnothing\}$, where the special symbol $\varnothing$ represents no output. Hence, if the source transmits and the relay receives, the source's and relay's input alphabets are $\mathcal{X}_1$ and $\{\varnothing\}$, respectively, whereas the relay's and destination's output alphabets are $\mathcal{Y}_1$ and $\{\varnothing\}$, respectively. On the other hand, if the relay transmits and the destination receives, the source's and relay's input alphabets are $\{\varnothing\}$ and $\mathcal{X}_2$, respectively, whereas the relay's and destination's output alphabets are $\{\varnothing\}$ and $\mathcal{Y}_2$, respectively.

We want to transmit message W, drawn uniformly from the message set $\{1, 2, ..., 2^{nR}\}$, from the source through the relay to the destination in n channel uses, for the case when all three

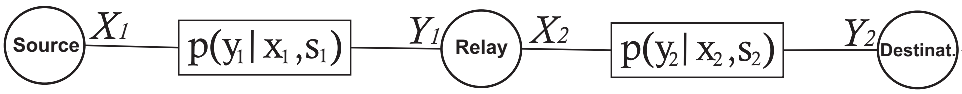

Fig. 2. System model comprising a source, a half-duplex relay equipped with a buffer, and a destination.

nodes have full knowledge of the channel states. To this end, source and relay transmit ordered sequences of symbols represented by random variables (RVs) $X_1^n = (X_{11}, X_{12}, ..., X_{1n})$ and $X_2^n = (X_{21}, X_{22}, ..., X_{2n})$, respectively, and relay and destination receive ordered sequences of symbols represented by RVs $Y_1^n = (Y_{11}, Y_{12}, ..., Y_{1n})$ and $Y_2^n = (Y_{21}, Y_{22}, ..., Y_{2n})$, respectively. During this transmission, the channel states on the source-relay and relay-destination channels are represented by RVs $S_1^n = (S_{11}, S_{12}, ..., S_{1n})$ and $S_2^n = (S_{21}, S_{22}, ..., S_{2n})$, respectively. For $i = 1, ..., n$, RVs X_{1i}, X_{2i}, Y_{1i}, Y_{2i}, S_{1i}, and S_{2i}, assume values denoted by x_{1i}, x_{2i}, y_{1i}, y_{2i}, s_{1i}, and s_{2i}, respectively, from the sets $\mathcal{X}_1^*$, $\mathcal{X}_2^*$, $\mathcal{Y}_1^*$, $\mathcal{Y}_2^*$, $\mathcal{S}_1$, and $\mathcal{S}_2$, respectively. For any channel use, the inputs at source and relay are represented by RVs X_1 and X_2, respectively, the outputs at the relay and the destination are represented by RVs Y_1 and Y_2, respectively, and the channel states on the source-relay and relay-destination channels are represented by RVs S_1 and S_2, respectively. RVs X_1, X_2, Y_1, Y_2, S_1, and S_2 take values denoted by x_1, x_2, y_1, y_2, s_1, and s_2, respectively, from the sets $\mathcal{X}_1^*$, $\mathcal{X}_2^*$, $\mathcal{Y}_1^*$, $\mathcal{Y}_2^*$, $\mathcal{S}_1$, and $\mathcal{S}_2$, respectively.

In the considered system, we assume causal CSI at all three nodes. More precisely, right before the i-th channel use, the source knows s_{1i}, the relay knows s_{1i} and s_{2i}, and the destination knows s_{2i} (we will prove that this is sufficient for achieving the capacity, cf. Section IV-B). Furthermore, the channel is memoryless in the sense that given the states and the input symbols for the i-th channel use, the i-th output symbols are independent from all previous states and input symbols, i.e.,

$$p(y_{1i}, y_{2i}|x_1^i, x_2^i, s_1^i, s_2^i) = p(y_{1i}, y_{2i}|x_{1i}, x_{2i}, s_{1i}, s_{2i}) \ , \tag{1}$$

where the notation a_j^i is used for the ordered sequence $a_j^i = (a_{j1}, a_{j2}, ..., a_{ji})$, where $a \in \{x, s\}$ and $j \in \{1, 2\}$. Moreover, since in the considered half-duplex relay channel the source-destination link is unavailable, the conditional PMF of the channel given by (1), can be further simplified as

$$p(y_{1i}, y_{2i}|x_{1i}, x_{2i}, s_{1i}, s_{2i}) = p(y_{1i}|x_{1i}, s_{1i})p(y_{2i}|x_{2i}, s_{2i}). \tag{2}$$

This means that for the i-th channel use, given the source input symbol x_{1i} and the source-relay channel state s_{1i}, the received symbol at the relay y_{1i} is independent of x_{2i}, y_{2i}, and s_{2i}, i.e., independent of the relay-destination channel. Similarly, for the i-th channel use, given the relay input symbol x_{2i} and the relay-destination channel state s_{2i}, the received symbol at the destination y_{2i} is independent of x_{1i}, y_{1i}, and s_{1i}, i.e., independent of the source-relay channel. As a direct consequence of (2), the system model in Fig. 1 can be represented by the equivalent system model in Fig. 2.

Finally, $I(\cdot\ ;\ \cdot)$ and $H(\cdot)$ denote the mutual information and the entropy, respectively.

III. The Channel Capacity

In this section, we provide the general capacity theorem and determine the capacity of three specific channels.

A. Capacity Theorem

Before we introduce the capacity expression, we first define all variables and probabilities that will be used in the capacity theorem. We represent the half-duplex constraint by the binary variable $d(s_1, s_2) \in \{0, 1\}$, such that $d(s_1, s_2) = 0$ if for states s_1 and s_2 the source input symbol is $X_1 \neq \varnothing$ and the relay input symbol is $X_2 = \varnothing$, i.e., the source transmits and the relay is silent (it receives) for states s_1 and s_2, whereas $d(s_1, s_2) = 1$ if for states s_1 and s_2 the source input symbol is $X_1 = \varnothing$ and the relay input symbol is $X_2 \neq \varnothing$, i.e., the source is silent and the relay transmits for states s_1 and s_2. Let $p(d(s_1, s_2) = 1|s_1, s_2)$ denote the probability that, given states s_1 and s_2, $d(s_1, s_2) = 1$, and let $p(d(s_1, s_2) = 0|s_1, s_2)$ denote the probability that, given states s_1 and s_2, $d(s_1, s_2) = 0$. Obviously, $p(d(s_1, s_2) = 0|s_1, s_2) + p(d(s_1, s_2) = 1|s_1, s_2) = 1$. Furthermore,

$$p(d(s_1, s_2) = 0, s_1, s_2) = p(d(s_1, s_2) = 0|s_1, s_2)p(s_1, s_2) \tag{3}$$

is the probability that states s_1 and s_2 occur and that $d(s_1, s_2) = 0$, whereas

$$p(d(s_1, s_2) = 1, s_1, s_2) = p(d(s_1, s_2) = 1|s_1, s_2)p(s_1, s_2) \tag{4}$$

is the probability that states s_1 and s_2 occur and that $d(s_1, s_2) = 1$. Furthermore, using $d(s_1, s_2) \in \{0, 1\}$ we note that conditioning X_1 and Y_1 on $d(s_1, s_2) = 0$ means that the source transmits

and the relay is silent, and therefore X_1 and Y_1 can take values only from sets $\mathcal{X}_1$ and $\mathcal{Y}_1$, respectively, and $X_1 = \varnothing$ and $Y_1 = \varnothing$ are not possible. Similarly, conditioning X_2 and Y_2 on $d(s_1, s_2) = 1$ means that the source is silent and the relay transmits, and therefore X_2 and Y_2 can take values only from sets $\mathcal{X}_2$ and $\mathcal{Y}_2$, respectively, and $X_2 = \varnothing$ and $Y_2 = \varnothing$ are not possible. Finally, as will be shown in the following, the optimal link selection policy may require a coin flip. For this purpose, we introduce the set of possible outcomes of the coin flip, $\mathcal{C} \in \{0, 1\}$, and denote the probabilities of the outcomes by $P_{\mathcal{C}} = \Pr\{\mathcal{C} = 1\}$ and $\Pr\{\mathcal{C} = 0\} = 1 - P_{\mathcal{C}}$, respectively. Now, we are ready to present the capacity of the state-dependent half-duplex relay channel without a source-destination link.

Theorem 1: Define variable $d(s_1, s_2)$ as

$$d(s_1,s_2) \triangleq \begin{cases} 1, \text{ if} & \max\limits_{p(x_2|d(s_1,s_2)=1,s_2)} I(X_2;Y_2|d(s_1,s_2)=1,s_2) \\ & > \rho \times \max\limits_{p(x_1|d(s_1,s_2)=0,s_1)} I(X_1;Y_1|d(s_1,s_2)=0,s_1) \\ \\ 0, \text{ if} & \max\limits_{p(x_2|d(s_1,s_2)=1,s_2)} I(X_2;Y_2|d(s_1,s_2)=1,s_2) \\ & < \rho \times \max\limits_{p(x_1|d(s_1,s_2)=0,s_1)} I(X_1;Y_1|d(s_1,s_2)=0,s_1) \\ \\ 1, \text{ if} & \max\limits_{p(x_2|d(s_1,s_2)=1,s_2)} I(X_2;Y_2|d(s_1,s_2)=1,s_2) \\ & = \rho \times \max\limits_{p(x_1|d(s_1,s_2)=0,s_1)} I(X_1;Y_1|d(s_1,s_2)=0,s_1) \text{ AND } \mathcal{C}=1 \\ \\ 0, \text{ if} & \max\limits_{p(x_2|d(s_1,s_2)=1,s_2)} I(X_2;Y_2|d(s_1,s_2)=1,s_2) \\ & = \rho \times \max\limits_{p(x_1|d(s_1,s_2)=0,s_1)} I(X_1;Y_1|d(s_1,s_2)=0,s_1) \text{ AND } \mathcal{C}=0, \end{cases} \tag{5}$$

where ρ is a constant and its value, as well as the value of $P_{\mathcal{C}}$, is determined as follows. Define $C_1(\rho, P_{\mathcal{C}})$ and $C_2(\rho, P_{\mathcal{C}})$ as

$$C_1\big(\rho, P_{\mathcal{C}}\big) \triangleq \sum_{s_1\in\mathcal{S}_1}\sum_{s_2\in\mathcal{S}_2} p(d(s_1,s_2)=0,s_1,s_2) \max_{p(x_1|d(s_1,s_2)=0,s_1)} I(X_1;Y_1|d(s_1,s_2)=0,s_1) \tag{6}$$

$$C_2\big(\rho, P_{\mathcal{C}}\big) \triangleq \sum_{s_1\in\mathcal{S}_1}\sum_{s_2\in\mathcal{S}_2} p(d(s_1,s_2)=1,s_1,s_2) \max_{p(x_2|d(s_1,s_2)=1,s_2)} I(X_2;Y_2|d(s_1,s_2)=1,s_2). \tag{7}$$

The optimal value of ρ, $\rho = \rho_{\text{opt}}$, and the optimal value of the coin flip probability $P_{\mathcal{C}}$, $P_{\mathcal{C}} = P_{\mathcal{C},\text{opt}}$, are the ones for which $C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ is achieved. Then, the capacity of the

considered state-dependent half-duplex relay channel, C, is given by

$$C = C_1\big(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}\big) = C_2\big(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}\big). \tag{8}$$

Proof: The proof of the converse and the achievability of the capacity are given in Sections IV-A and IV-B, respectively. ■

The interpretation of $C_1(\rho, P_{\mathcal{C}})$ and $C_2(\rho, P_{\mathcal{C}})$ is that $C_1(\rho, P_{\mathcal{C}})$ is the average rate achieved in the source-relay channel for a given ρ and $P_{\mathcal{C}}$, and $C_2(\rho, P_{\mathcal{C}})$ is the average rate achieved in the relay-destination channel for a given ρ and $P_{\mathcal{C}}$, under the assumption that the relay always has enough information to transmit. The capacity is obtained when $C_1(\rho_{\rm opt}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\rm opt}, P_{\mathcal{C},\text{opt}})$ is achieved for some $\rho = \rho_{\rm opt}$ and $P_{\mathcal{C}} = P_{\mathcal{C},\text{opt}}$, i.e., when the amount of information that the source transmits is equal to the amount of information that the relay transmits. Furthermore, if $C_1(\rho_{\rm opt}, P_{\mathcal{C}}) = C_2(\rho_{\rm opt}, P_{\mathcal{C}})$ is achieved independent of $P_{\mathcal{C}}$, then this means that the coin flip is not used, i.e., for every $s_1 \in \mathcal{S}_1$ and $s_2 \in \mathcal{S}_2$, the values of $d(s_1, s_2)$ are determined only be the first two conditions in (5), and not by the last two. On the other hand, if $C_1(\rho_{\rm opt}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\rm opt}, P_{\mathcal{C},\text{opt}})$ is achieved by using the coin flip, then for those $s_1 \in \mathcal{S}_1$ and $s_2 \in \mathcal{S}_2$, for which a coin flip is used, the value of $d(s_1, s_2)$ is identical to the outcome of the coin flip $\mathcal{C}$.

B. Application of the Capacity Theorem

In the following, we consider three specific channels for which we determine the capacity.

1) Fixed Channel: Let both sets $\mathcal{S}_1$ and $\mathcal{S}_2$ be comprised of only one element, i.e., $\mathcal{S}_1 = \{s_1\}$ and $\mathcal{S}_2 = \{s_2\}$. This models a half-duplex relay channel with one state only which includes for example the case when both the source-relay and the relay-destination link are non-fading AWGN[1] channels. For s_1 and s_2 let us denote

$$\max_{p(x_1|d(s_1,s_2)=0,s_1)} I(X_1;Y_1|d(s_1,s_2)=0,s_1) \quad = \quad A \tag{9}$$

$$\max_{p(x_2|d(s_1,s_2)=1,s_2)} I(X_2;Y_2|d(s_1,s_2)=1,s_2) \quad = \quad B, \tag{10}$$

where $A > 0$ and $B > 0$. This is an example in which the condition $C_1\big(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}\big) = C_2\big(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}\big)$ can only be achieved if ρ is set to be such that the first two conditions in

[1]It is well known that capacity theorems for channels with finite input and finite output alphabets can be straightforwardly extended to AWGN channels [7].

(5) never occur and only the last two occur. Thus, a coin flip has to be used and the value of $d(s_1, s_2)$ is identical to the outcome of the coin flip. More precisely, ρ_{opt} that achieves $C_1\big(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}\big) = C_2\big(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}\big)$ is given by

$$\rho_{\text{opt}} = \frac{B}{A}. \tag{11}$$

It is straightforward to show that for any $\rho \neq \rho_{\text{opt}}$, $C_1(\rho, P_\mathcal{C}) = C_2(\rho, P_\mathcal{C})$ cannot be achieved. In particular, if ρ is chosen such that $\rho < \rho_{\text{opt}}$, then it can be seen from (5) that $d(s_1, s_2)$ will always be $d(s_1, s_2) = 1$. Thereby, $C_1(\rho, P_\mathcal{C})$ and $C_2(\rho, P_\mathcal{C})$, given by (6) and (7), respectively, become

$$C_1(\rho, P_\mathcal{C}) \;=\; 0 \tag{12}$$

$$C_2(\rho, P_\mathcal{C}) \;=\; B. \tag{13}$$

Hence, for $\rho < \rho_{\text{opt}}$, $C_1(\rho, P_\mathcal{C}) = C_2(\rho, P_\mathcal{C})$ cannot be achieved. Similarly, if $\rho > \rho_{\text{opt}}$, $d(s_1, s_2)$ will always be $d(s_1, s_2) = 0$, and $C_1(\rho, P_\mathcal{C})$ and $C_2(\rho, P_\mathcal{C})$ become

$$C_1(\rho, P_\mathcal{C}) \;=\; A \tag{14}$$

$$C_2(\rho, P_\mathcal{C}) \;=\; 0. \tag{15}$$

Hence, for $\rho > \rho_{\text{opt}}$, $C_1(\rho, P_\mathcal{C}) = C_2(\rho, P_\mathcal{C})$ also cannot be achieved. Therefore, the only value left is $\rho = \rho_{\text{opt}}$. For $\rho = \rho_{\text{opt}}$, we now have the outcome of the coin flip $\mathcal{C} \in \{0, 1\}$ as an additional parameter which decides the value of $d(s_1, s_2)$. In particular, for $\rho = \rho_{\text{opt}}$, $d(s_1, s_2) = 0$ if $\mathcal{C} = 0$ and this happens with probability $1 - P_\mathcal{C}$, and $d(s_1, s_2) = 1$ if $\mathcal{C} = 1$ and this happens with probability $P_\mathcal{C}$. Therefore, $C_1(\rho_{\text{opt}}, P_\mathcal{C})$ and $C_2(\rho_{\text{opt}}, P_\mathcal{C})$ are given by

$$C_1(\rho_{\text{opt}}, P_\mathcal{C}) \;=\; (1 - P_\mathcal{C}) \times A \tag{16}$$

$$C_2(\rho_{\text{opt}}, P_\mathcal{C}) \;=\; P_\mathcal{C} \times B. \tag{17}$$

From (16) and (17), we obtain $P_{\mathcal{C},\text{opt}}$ as the $P_\mathcal{C}$ for which $C_1(\rho_{\text{opt}}, P_\mathcal{C}) = C_2(\rho_{\text{opt}}, P_\mathcal{C})$ holds, i.e.,

$$P_{\mathcal{C},\text{opt}} = \frac{A}{A + B}. \tag{18}$$

Inserting (18) into (16) or (17), we obtain the capacity $C = C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ as

$$C = \frac{AB}{A + B}. \tag{19}$$

Remark 1: Of course, as expected, the capacity that we have obtained in this example is exactly the one already obtained in the literature for the AWGN channel without fading, see [2] for example. We note however that there is a difference from the previous result in the time sharing approach. In particular, we use random time sharing via the coin flip as opposed to the deterministic time sharing in [2]. The random time sharing approach is more general and allows us to obtain the capacity of a large class of state-dependent half-duplex relay channels.

2) ON-OFF Channel: In the second example, the sets $\mathcal{S}_1$ and $\mathcal{S}_2$ are both comprised of two elements given by $\mathcal{S}_1 = \{s_1[1], s_1[2]\}$ and $\mathcal{S}_2 = \{s_2[1], s_2[2]\}$. Let the channel states be such that

$$\max_{p(x_1|d(s_1,s_2)=0,s_1[1])} I(X_1;Y_1|d(s_1,s_2)=0,s_1[1]) = 0 \tag{20}$$

$$\max_{p(x_2|d(s_1,s_2)=1,s_2[1])} I(X_2;Y_2|d(s_1,s_2)=1,s_2[1]) = 0 \tag{21}$$

$$\max_{p(x_1|d(s_1,s_2)=0,s_1[2])} I(X_1;Y_1|d(s_1,s_2)=0,s_1[2]) = A \tag{22}$$

$$\max_{p(x_2|d(s_1,s_2)=1,s_2[2])} I(X_2;Y_2|d(s_1,s_2)=1,s_2[2]) = B, \tag{23}$$

where A and B are constants larger than zero. In other words, the source-relay and relay-destination channels are ON-OFF channels such that for states $s_1 = s_1[1]$ and $s_2 = s_2[1]$ the source-relay and relay-destination channels, respectively, are OFF, i.e., both have zero capacity, and for states $s_1 = s_1[2]$ and $s_2 = s_2[2]$ the source-relay and relay-destination channels, respectively, are ON, i.e., the source-relay and relay-destination channels have capacities A and B, respectively. Furthermore, since for $s_1 = s_1[1]$ the capacity of the source-relay channel is zero and since the source knows s_1, the source will be silent if $s_1 = s_1[1]$ occurs, i.e., it can achieve zero rate by not transmitting. Similar, since for $s_2 = s_2[1]$ the capacity of the relay-destination channel is zero and since the relay knows s_2, the relay will be silent if $s_2 = s_2[1]$ occurs, i.e., it can achieve zero rate by not transmitting.

In this example, depending on the values of A and B, and the statistics of S_1 and S_2, three cases have to be distinguished. In all three cases, the condition $C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ can only be achieved if a coin flip is used. However, each of the three cases results in different values for the pair $(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$, since the condition $C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ cannot be satisfied by using only one value for the pair $(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ for all values of A and B and all statistics of S_1 and S_2. Before discussing the three cases, we note that the maximal average rate

that the source (relay) can achieve in the source-relay (relay-destination) channel is obtained when the source (relay) transmits with rate A (B) in all occurrences of state $s_1[2]$ ($s_2[2]$), and this average rate is $p(s_1[2])A$ ($p(s_2[2])B$). The three cases are provided in the following.

Case 1) The first case is valid when the constants A and B, and the RVs S_1 and S_2 are such that

$$p(s_2[2])B < p(s_1[2], s_2[1])A \tag{24}$$

holds. In other words, in this case the relay transmits with rate B if state $s_2[2]$ occurs and the maximum achievable average rate of the relay-destination link, $p(s_2[2])B$, is smaller than the average rate $p(s_1[2], s_2[1])A$ obtained when the source always transmits when the relay is silent. In this case, the condition $C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ can only be achieved for $\rho_{\text{opt}} = 0$, i.e., if ρ is set to any other value $\rho > \rho_{\text{opt}} = 0$, $C_1(\rho, P_{\mathcal{C}}) = C_2(\rho, P_{\mathcal{C}})$ cannot hold. To show this, we insert (20)-(23) into (5) and set $\rho > \rho_{\text{opt}} = 0$. Then, we note that for any such $\rho > \rho_{\text{opt}} = 0$, $C_1(\rho, P_{\mathcal{C}}) \geq p(s_1[2], s_2[1])A$ and $C_2(\rho, P_{\mathcal{C}}) \leq p(s_2[2])B$ hold. Hence, using these two inequalities together with (24), we obtain that $C_1(\rho, P_{\mathcal{C}}) > C_2(\rho, P_{\mathcal{C}})$ holds always for $\rho > \rho_{\text{opt}} = 0$. Therefore, the only value left is $\rho = \rho_{\text{opt}} = 0$ and this will require usage of the coin flip. By inserting (20)-(23) into (5) and setting $\rho = 0$, we obtain the values of $d(s_1, s_2)$ for all four combinations of $(s_1, s_2) = \{(s_1[1], s_2[1]), (s_1[1], s_2[2]), (s_1[2], s_2[1]), (s_1[2], s_2[2])\}$. In particular, since for $(s_1, s_2) = (s_1[1], s_2[1])$ both nodes are silent, i.e., transmit with zero rate, the value if $d(s_1, s_2)$ is not important. For $(s_1, s_2) = (s_1[1], s_2[2])$, we obtain

$$d(s_1, s_2) = \begin{cases} 1, & \text{if} \quad B > 0 \times 0 \\ 0, & \text{if} \quad B < 0 \times 0 \\ 1, & \text{if} \quad B = 0 \times 0 \text{ AND } \mathcal{C} = 1 \\ 0, & \text{if} \quad B = 0 \times 0 \text{ AND } \mathcal{C} = 0. \end{cases} \tag{25}$$

Hence, $d(s_1, s_2) = 1$ for $(s_1, s_2) = (s_1[1], s_2[2])$. For $(s_1, s_2) = (s_1[2], s_2[2])$, we obtain

$$d(s_1, s_2) = \begin{cases} 1, & \text{if} \quad B > 0 \times A \\ 0, & \text{if} \quad B < 0 \times A \\ 1, & \text{if} \quad B = 0 \times A \text{ AND } \mathcal{C} = 1 \\ 0, & \text{if} \quad B = 0 \times A \text{ AND } \mathcal{C} = 0. \end{cases} \tag{26}$$

Hence, $d(s_1, s_2) = 1$ for $(s_1, s_2) = (s_1[2], s_2[2])$. Finally, for $(s_1, s_2) = (s_1[2], s_2[1])$, we obtain

$$d(s_1, s_2) = \begin{cases} 1, & \text{if} \quad 0 > 0 \times A \\ 0, & \text{if} \quad 0 < 0 \times A \\ 1, & \text{if} \quad 0 = 0 \times A \text{ AND } \mathcal{C} = 1 \\ 0, & \text{if} \quad 0 = 0 \times A \text{ AND } \mathcal{C} = 0. \end{cases} \tag{27}$$

Hence, the value of $d(s_1, s_2)$ is identical to the outcome of the coin flip, i.e., for $(s_1, s_2) = (s_1[2], s_2[1])$, $d(s_1, s_2) = 1$ if $\mathcal{C} = 1$ occurs and $d(s_1, s_2) = 0$ if $\mathcal{C} = 0$ occurs. Thereby, the source transmits with rate A only if $(s_1, s_2) = (s_1[2], s_2[1])$ and $\mathcal{C} = 0$ occur, whereas the relay transmits with rate B in all occurrences of $s_2 = s_2[2]$. Hence, for $\rho = \rho_{\text{opt}} = 0$ and arbitrary $P_{\mathcal{C}}$, $C_1(\rho_{\text{opt}}, P_{\mathcal{C}})$ and $C_1(\rho_{\text{opt}}, P_{\mathcal{C}})$ are obtained as

$$C_1(\rho_{\text{opt}}, P_{\mathcal{C}}) = (1 - P_{\mathcal{C}}) \times p(s_1[2], s_2[1])A \tag{28}$$

$$C_2(\rho_{\text{opt}}, P_{\mathcal{C}}) = p(s_2[2])B. \tag{29}$$

From (28) and (29) we obtain $P_{\mathcal{C},\text{opt}}$ as the $P_{\mathcal{C}}$ for which $C_1(\rho_{\text{opt}}, P_{\mathcal{C}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C}})$ is achieved, i.e.,

$$P_{\mathcal{C},\text{opt}} = 1 - \frac{p(s_2[2])B}{p(s_1[2], s_2[1])A}. \tag{30}$$

Inserting (30) into (28), the capacity, $C = C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$, is obtained as

$$C = p(s_2[2])B. \tag{31}$$

Note that this is also the average capacity of the relay-destination channel.

Case 2) The second case is identical to Case 1) with source-relay and relay-destination channels switching places. In particular, this case is valid when the constants A and B, and the RVs S_1 and S_2 are such that

$$p(s_1[2])A < p(s_1[1], s_2[2])B \tag{32}$$

holds. In other words, this case is valid when the source transmits with rate A in all occurrences of state $s_1[2]$ and yet the maximum achievable average rate by the source, $p(s_1[2])A$, is smaller than the average rate $p(s_1[1], s_2[2])B$ obtained when the relay transmits always when the source

is silent. Following a similar procedure as in Case 1), we obtain that in this case $\rho_{\text{opt}} = \infty$, and $C_1(\rho_{\text{opt}}, P_\mathcal{C})$ and $C_2(\rho_{\text{opt}}, P_\mathcal{C})$ are obtained as

$$C_1(\rho_{\text{opt}}, P_\mathcal{C}) = p(s_1[2])A \tag{33}$$

$$C_2(\rho_{\text{opt}}, P_\mathcal{C}) = P_\mathcal{C} \times p(s_1[1], s_2[2])B. \tag{34}$$

From (33) and (34), we obtain $P_{\mathcal{C},\text{opt}}$ as the $P_\mathcal{C}$ for which $C_1(\rho_{\text{opt}}, P_\mathcal{C}) = C_2(\rho_{\text{opt}}, P_\mathcal{C})$ is achieved, i.e.,

$$P_{\mathcal{C},\text{opt}} = \frac{p(s_1[2])A}{p(s_1[1], s_2[2])B}. \tag{35}$$

Thereby, the capacity, $C = C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$, is obtained as

$$C = p(s_1[2])A. \tag{36}$$

Note that this is also the average capacity of the source-relay channel.

Case 3) The third and final case is when the constants A and B, and the RVs S_1 and S_2 are such that both (24) and (32) do not hold. In this case, the condition $C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ can be achieved only if ρ is set to be $\rho = \rho_{\text{opt}} = B/A$. Otherwise, if $\rho < \rho_{\text{opt}} = B/A$, then $d(s_1, s_2) = 1$, $\forall (s_1, s_2)$, and therefore $C_1(\rho, P_\mathcal{C}) = 0$ and $C_2(\rho, P_\mathcal{C}) = p(s_2[2])B$. Whereas, if $\rho > \rho_{\text{opt}} = B/A$, then $d(s_1, s_2) = 0$, $\forall (s_1, s_2)$, and therefore $C_1(\rho, P_\mathcal{C}) = p(s_1[2])A$ and $C_2(\rho, P_\mathcal{C}) = 0$. By inserting $\rho = \rho_{\text{opt}} = B/A$ and (20)-(23) into (5), we obtain the values of $d(s_1, s_2)$ for all combinations of $(s_1, s_2) = \{(s_1[1], s_2[1]), (s_1[1], s_2[2]), (s_1[2], s_2[1]), (s_1[2], s_2[2])\}$. Since for $(s_1, s_2) = (s_1[1], s_2[1])$ both source and relay are silent, the value of $d(s_1, s_2)$ is not important. For $(s_1, s_2) = (s_1[1], s_2[2])$, we obtain

$$d(s_1, s_2) = \begin{cases} 1, & \text{if} \quad B > B/A \times 0 \\ 0, & \text{if} \quad B < B/A \times 0 \\ 1, & \text{if} \quad B = B/A \times 0 \text{ AND } \mathcal{C} = 1 \\ 0, & \text{if} \quad B = B/A \times 0 \text{ AND } \mathcal{C} = 0. \end{cases} \tag{37}$$

Hence, $d(s_1, s_2) = 1$ for $(s_1, s_2) = (s_1[1], s_2[2])$. For $(s_1, s_2) = (s_1[2], s_2[1])$, we obtain

$$d(s_1, s_2) = \begin{cases} 1, & \text{if} \quad 0 > B/A \times A \\ 0, & \text{if} \quad 0 < B/A \times A \\ 1, & \text{if} \quad 0 = B/A \times A \text{ AND } \mathcal{C} = 1 \\ 0, & \text{if} \quad 0 = B/A \times A \text{ AND } \mathcal{C} = 0. \end{cases} \tag{38}$$

Hence, $d(s_1, s_2) = 0$ for $(s_1, s_2) = (s_1[2], s_2[1])$. Finally, for $(s_1, s_2) = (s_1[2], s_2[2])$, we obtain

$$d(s_1, s_2) = \begin{cases} 1, & \text{if} \quad B > B/A \times A \\ 0, & \text{if} \quad B < B/A \times A \\ 1, & \text{if} \quad B = B/A \times A \text{ AND } \mathcal{C} = 1 \\ 0, & \text{if} \quad B = B/A \times A \text{ AND } \mathcal{C} = 0. \end{cases} \tag{39}$$

Hence, the coin flip is used and $d(s_1, s_2) = 1$ if $\mathcal{C} = 1$ and $d(s_1, s_2) = 0$ if $\mathcal{C} = 0$. Thereby, the source transmits with rate A if $(s_1[2], s_2[1])$ occurs or if $(s_1[2], s_2[2])$ and $\mathcal{C} = 0$ occur, whereas the relay transmits with rate B if $(s_1[1], s_2[2])$ occurs or if $(s_1[2], s_2[2])$ and $\mathcal{C} = 1$ occur. Hence, for $\rho = \rho_{\text{opt}} = B/A$ and arbitrary $P_{\mathcal{C}}$, $C_1(\rho_{\text{opt}}, P_{\mathcal{C}})$ and $C_2(\rho_{\text{opt}}, P_{\mathcal{C}})$ are obtained as

$$C_1(\rho_{\text{opt}}, P_{\mathcal{C}}) = p(s_1[2], s_2[1])A + (1 - P_{\mathcal{C}}) \times p(s_1[2], s_2[2])A \tag{40}$$

$$C_2(\rho_{\text{opt}}, P_{\mathcal{C}}) = p(s_1[1], s_2[2])B + P_{\mathcal{C}} \times p(s_1[2], s_2[2])B. \tag{41}$$

From (40) and (41), $P_{\mathcal{C},\text{opt}}$ is found as the $P_{\mathcal{C}}$ for which $C_1(\rho_{\text{opt}}, P_{\mathcal{C}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C}})$ is achieved. Then inserting $P_{\mathcal{C},\text{opt}}$ in (40) or (41) the capacity $C = C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ is obtained as

$$C = \frac{AB}{A + B}\Big(1 - p(s_1[1], s_2[1])\Big). \tag{42}$$

Note that for $p(s_1[1], s_2[1]) = 0$, this is identical to the capacity given by (19) for fixed channels.

3) Fading AWGN Channel: In the third example, we make certain assumptions and approximations so that Theorem 1, which was derived for finite sets $\mathcal{S}_1$ and $\mathcal{S}_2$, is applicable to the case when sets $\mathcal{S}_1$ and $\mathcal{S}_2$ contain infinite numbers of states. In particular, let the states in the sets $\mathcal{S}_1$ and $\mathcal{S}_2$ be arranged in ascending order as $\mathcal{S}_1 = \{s_1[1], s_1[2], ..., s_1[M_1]\}$ and $\mathcal{S}_2 = \{s_2[1], s_2[2], ..., s_2[M_2]\}$, respectively, where $M_1 \to \infty$ and $M_2 \to \infty$. Moreover, the difference between $s_1[i]$ and $s_1[i+1]$ for any $i = 1, ..., M_1 - 1$, and the difference between $s_2[i]$ and $s_2[i+1]$ for any $i = 1, ..., M_2 - 1$, are such that $s_1[i] - s_1[i+1] \to 0$ when $M_1 \to \infty$ and $s_2[i] - s_2[i+1] \to 0$ when $M_2 \to \infty$. In other words, the values in the ordered sets $\mathcal{S}_1 = \{s_1[1], s_1[2], ..., s_1[M_1]\}$ and $\mathcal{S}_2 = \{s_2[1], s_2[2], ..., s_2[M_1]\}$ form a continuous line for $M_1 \to \infty$ and $M_2 \to \infty$. A similar continuity holds for the mutual informations of the source-relay channel and relay-destination channel, respectively. In particular, for the mutual

informations the following holds

$$I(X_1;Y_1|d(s_1,s_2)=0,s_1[i])-I(X_1;Y_1|d(s_1,s_2)=0,s_1[i+1])\to 0,\ \text{as } M_1\to\infty, \tag{43}$$

$$I(X_2;Y_2|d(s_1,s_2)=1,s_2[i])-I(X_2;Y_2|d(s_1,s_2)=1,s_2[i+1])\to 0,\ \text{as } M_2\to\infty, \tag{44}$$

i.e., the mutual informations are smooth functions without discontinuities. This models the case when the source-relay and relay-destination channels are impaired by time-continuous fading, e.g., wireless Rayleigh fading channels.

If the above assumptions are satisfied, the probability that states (s_1,s_2) occur such that

$$\max_{p(x_2|d(s_1,s_2)=1,s_2)} I(X_2;Y_2|d(s_1,s_2)=1,s_2)=\rho\times\max_{p(x_1|d(s_1,s_2)=0,s_1)} I(X_1;Y_1|d(s_1,s_2)=0,s_1) \tag{45}$$

holds, for any constant ρ, is negligible compared to the probability for which (45) does not hold. Therefore, the probability that $d(s_1,s_2)$ in (5) requires a coin flip is negligible compared to the probability that a coin flip is not required, i.e., practically only the first two conditions in (5) will occur and not the last two. As a result, $C_1(\rho,P_{\mathcal{C}})=C_1(\rho)$ and $C_2(\rho,P_{\mathcal{C}})=C_2(\rho)$. Furthermore, we can replace the sums in (6) and (7) by integrals and obtain

$$\begin{aligned}C_1(\rho)=\int_{s_1=0}^{\infty}\int_{s_2=0}^{\infty}\max_{p(x_1|d(s_1,s_2)=0,s_1)}&I(X_1;Y_1|d(s_1,s_2)=0,s_1)\\&\times f(d(s_1,s_2)=0,s_1,s_2)ds_1ds_2\end{aligned} \tag{46}$$

$$\begin{aligned}C_2(\rho)=\int_{s_1=0}^{\infty}\int_{s_2=0}^{\infty}\max_{p(x_2|d(s_1,s_2)=1,s_2)}&I(X_2;Y_2|d(s_1,s_2)=1,s_2)\\&\times f(d(s_1,s_2)=1,s_1,s_2)ds_1ds_2,\end{aligned} \tag{47}$$

where the PMFs $p(d(s_1,s_2)=0,s_1,s_2)$ and $p(d(s_1,s_2)=1,s_1,s_2)$ are replaced by probability distribution functions (PDFs) $f(d(s_1,s_2)=0,s_1,s_2)$ and $f(d(s_1,s_2)=1,s_1,s_2)$, respectively, where

$$\int_{s_1=a_1}^{b_1}\int_{s_2=a_2}^{b_2} f(d(s_1,s_2)=0,s_1,s_2)ds_1ds_2$$

represents the probability (or fraction of the total time) that states s_1 and s_2 will take values in the domain $[a_1,b_1]$ and $[a_2,b_2]$, respectively, and $d(s_1,s_2)=0$. A similar interpretation holds for $f(d(s_1,s_2)=1,s_1,s_2)$.

If the source-relay and relay-destination channels are AWGN channels with time-continuous fading, such that when state s_1 occurs the SNR on the source-relay channel is γ_1, whereas when

state s_2 occurs the SNR on the relay-destination channel is γ_2, then

$$\max_{p(x_1|d(s_1,s_2)=0,s_1)} I(X_1;Y_1|d(s_1,s_2)=0,s_1) = \log_2(1+\gamma_1) \quad (48)$$

$$\max_{p(x_2|d(s,s_2)=1,s_2)} I(X_2;Y_2|d(s_1,s_2)=1,s_2) = \log_2(1+\gamma_2) \quad (49)$$

$$d(s_1,s_2) = d(\gamma_1,\gamma_2). \quad (50)$$

Inserting (48)-(50) into (5), we obtain that $d(\gamma_1,\gamma_2)=1$ if $\gamma_2 > (1+\gamma_1)^\rho - 1$, and $d(\gamma_1,\gamma_2)=0$ otherwise, i.e., if $\gamma_1 > (1+\gamma_2)^{1/\rho} - 1$. Hence, by inserting (48) and (49) into (46) and (47), respectively, $C_1(\rho)$ and $C_2(\rho)$ can be written as

$$\begin{aligned} C_1(\rho) &= \int_{\gamma_1=0}^{\infty}\int_{\gamma_2=0}^{\infty} \log_2(1+\gamma_1)\times f(d(\gamma_1,\gamma_2)=0,\gamma_1,\gamma_2)d\gamma_1 d\gamma_2 \\ &= \int_{\gamma_2=0}^{\infty}\int_{\gamma_1=(1+\gamma_2)^{1/\rho}-1}^{\infty} \log_2(1+\gamma_1)\times f(\gamma_1,\gamma_2)d\gamma_1 d\gamma_2 \quad (51) \\ C_2(\rho) &= \int_{\gamma_1=0}^{\infty}\int_{\gamma_2=0}^{\infty} \log_2(1+\gamma_2)\times f(d(\gamma_1,\gamma_2)=1,\gamma_1,\gamma_2)d\gamma_1 d\gamma_2 \\ &= \int_{\gamma_1=0}^{\infty}\int_{\gamma_2=(1+\gamma_1)^{\rho}-1}^{\infty} \log_2(1+\gamma_2)\times f(\gamma_1,\gamma_2)d\gamma_1 d\gamma_2, \quad (52) \end{aligned}$$

where $f(\gamma_1,\gamma_2)$ is the joint PDF of the SNRs γ_1 and γ_2. From (51) and (52), we can find the optimal ρ_{opt} for which $C_1(\rho_{\text{opt}}) = C_2(\rho_{\text{opt}})$ is achieved, and thereby obtain the capacity $C = C_1(\rho_{\text{opt}}) = C_2(\rho_{\text{opt}})$. As expected, this result is identical to the achievable rate given in [5], thereby the rate in [5] is the capacity.

IV. Converse and Achievability

In the following, we prove the converse and the achievability of the capacity, and therby prove Theorem 1.

A. Proof of Converse

Theorem 2: Any sequence of $(2^{nR}, n)$ codes with probability of error $P_e^{(n)} \to 0$ as $n \to \infty$, must have $R \le C$, where C is given by Theorem 1.

Proof: Let there be a $(2^{nR}, n)$ code with $P_e^{(n)} \to 0$ as $n \to \infty$. The source transmits message W drawn uniformly from the message set $\{1, 2, ..., 2^{nR}\}$, and the destination decodes

the received information to massage $\hat{W}$. Then, nR can be written as

$$
\begin{aligned}
&nR = H(W|s_1^n, s_2^n) = H(W|\hat{W}, s_1^n, s_2^n) + I(W;\hat{W}|s_1^n, s_2^n) \overset{(a)}{\leq} 1 + P_e^{(n)}nR + I(W;\hat{W}|s_1^n, s_2^n)\\
&\overset{(b)}{\leq} 1 + P_e^{(n)}nR + I(X_1^n;Y_2^n|s_1^n, s_2^n) \overset{(c)}{\leq} 1 + P_e^{(n)}nR + \min\left\{I\big(X_1^n;Y_1^n|X_2^n, s_1^n\big), I\big(X_2^n;Y_2^n|s_2^n\big)\right\}\\
&\overset{(d)}{\leq} 1 + P_e^{(n)}nR + \min\left\{\sum_{i=1}^n I\big(X_{1i};Y_{1i}|X_{2i}, s_{1i}\big), \sum_{i=1}^n I\big(X_{2i};Y_{2i}|s_{2i}\big)\right\}\\
&\overset{(e)}{\leq} 1 + P_e^{(n)}nR + \sup\min\left\{\sum_{i=1}^n I\big(X_{1i};Y_{1i}|X_{2i}, s_{1i}\big), \sum_{i=1}^n I\big(X_{2i};Y_{2i}|s_{2i}\big)\right\},
\end{aligned}
\tag{53}
$$

where (a), (b), (c), (d), and (e) follow from Fano's inequality [1, Sec. 2.11], the data processing inequality [1, Sec. 2.8], the cut-set bound [8, Sec. 1.2], the memoryless channel assumption [1, Sec. 7], and from taking the supremum over all variables with a degree of freedom for the expression in the $\min\{\cdot\}$ function, respectively. Furthermore, in (53) we have assumed codes constructed for the case when the channel states $s_1^n = (s_{11}, s_{12}, ..., s_{1n})$ and $s_2^n = (s_{21}, s_{12}, ..., s_{2n})$ are known a priori, since any code with a priori knowledge of the channel states can have a rate at least as high as the best code without a priori knowledge of the channel states. Due to the half-duplex transmission, the following holds

$$
I\big(X_{1i};Y_{1i}|X_{2i}, s_{1i}\big) = \begin{cases} I\big(X_{1i};Y_{1i}|X_{2i} = \varnothing, s_{1i}\big) \\ 0, \quad \text{if } X_{2i} \neq \varnothing \end{cases} \tag{54}
$$

$$
I\big(X_{2i};Y_{2i}|s_{2i}\big) = \begin{cases} I\big(X_{2i};Y_{2i}|X_{2i} \neq \varnothing, s_{2i}\big) \\ 0, \quad \text{if } X_{2i} = \varnothing \end{cases} = \begin{cases} I\big(X_{2i};Y_{2i}|X_{1i} = \varnothing, s_{2i}\big) \\ 0, \quad \text{if } X_{1i} \neq \varnothing, \end{cases} \tag{55}
$$

where the right hand side of (55) is transformed into a similar form as the right hand side of (54) by replacing $X_{2i} = \varnothing$ with $X_{1i} \neq \varnothing$, and $X_{2i} \neq \varnothing$ with $X_{1i} = \varnothing$. Hence, we can introduce a binary variable, $d_i \in \{0, 1\}$, such that $d_i = 0$ if $X_{1i} \neq \varnothing$ and $X_{2i} = \varnothing$, and $d_i = 1$ if $X_{1i} = \varnothing$ and $X_{2i} \neq \varnothing$. In other words, $d_i = 0$ when the source transmits and the relay is silent (it receives), and $d_i = 1$ when the relay transmits and the source is silent. Note that having $X_{1i} = X_{2i} = \varnothing$, i.e., both nodes are silent, will provide a rate which is smaller or equal to the rate when we exclude $X_{1i} = X_{2i} = \varnothing$. Since we are interested in an upper bound here, the case $X_{1i} = X_{2i} = \varnothing$ is not relevant and can be excluded. Using d_i, (54) and (55) can be written in compact form as

$$
I\big(X_{1i};Y_{1i}|X_{2i}, s_{1i}\big) = (1 - d_i)I\big(X_{1i};Y_{1i}|d_i = 0, s_{1i}\big), \tag{56}
$$

and

$$I\big(X_{2i};Y_{2i}|s_{2i}\big) = d_i I\big(X_{2i};Y_{2i}|d_i=1,s_{2i}\big), \tag{57}$$

respectively. Inserting (56) and (57) in (53), and dividing both sides by n, we obtain

$$R \leq \frac{1}{n} + P_e^{(n)} R + \frac{1}{n} \sup \min \left\{ \sum_{i=1}^{n} (1-d_i) I\big(X_{1i};Y_{1i}|d_i=0,s_{1i}\big) \,,\, \sum_{i=1}^{n} d_i I\big(X_{2i};Y_{2i}|d_i=1,s_{2i}\big) \right\}. \tag{58}$$

Since we assumed $P_e^{(n)} \to 0$ when $n \to \infty$, (58) can be simplified to

$$R \leq \frac{1}{n} \sup \min \left\{ \sum_{i=1}^{n} (1-d_i) I\big(X_{1i};Y_{1i}|d_i=0,s_{1i}\big) \,,\, \sum_{i=1}^{n} d_i I\big(X_{2i};Y_{2i}|d_i=1,s_{2i}\big) \right\}. \tag{59}$$

The only variables with a degree of freedom over which the supremum can be taken in (59) are $d_i \in \{0,1\}$, $p(x_{1i}|d_i=0,s_{1i})$, and $p(x_{2i}|d_i=1,s_{2i})$, for $i=1,...,n$. By inserting the maximization over $p(x_{1i}|d_i=0,s_{1i})$ and $p(x_{2i}|d_i=1,s_{2i})$ in (59), we are left only with the supremum over d_i, $\forall i$. Hence, (59) can be written as

$$\begin{aligned} R \leq \sup_{\{d_i\}_{i=1}^n} \min \Bigg\{ & \frac{1}{n} \sum_{i=1}^{n} (1-d_i) \max_{p(x_{1i}|d_i=0,s_{1i})} I\big(X_{1i};Y_{1i}|d_i=0,s_{1i}\big) \,, \\ & \frac{1}{n} \sum_{i=1}^{n} d_i \max_{p(x_{2i}|d_i=1,s_{2i})} I\big(X_{2i};Y_{2i}|d_i=1,s_{2i}\big) \Bigg\}, \end{aligned} \tag{60}$$

where it is assumed that the normalized sums in the $\min\{\cdot\}$ function in (60) have a limit for $n \to \infty$. The first term inside the $\min\{\cdot\}$ function in (60) increases as the number of d_i-s equal to zero increases. On the contrary, the second term inside the $\min\{\cdot\}$ function in (60) decreases as the number of d_i-s equal to zero increases. Therefore, the right hand side of (60) is maximized if and only if both terms in the $\min\{\cdot\}$ function become equal. The opposite cannot be true since one can always increase the smaller term at the expense of the larger term in the $\min\{\cdot\}$ function in (60). In other words, finding the supremum in (60) is equivalent to solving

$$\underset{\{d_i\}_{i=1}^n}{\text{Maximize}:} \quad \frac{1}{n} \sum_{i=1}^{n} (1-d_i) \max_{p(x_{1i}|d_i=0,s_{1i})} I\big(X_{1i};Y_{1i}|d_i=0,s_{1i}\big) \tag{61}$$

or

$$\underset{\{d_i\}_{i=1}^n}{\text{Maximize}:} \quad \frac{1}{n} \sum_{i=1}^{n} d_i \max_{p(x_{2i}|d_i=1,s_{2i})} I\big(X_{2i};Y_{2i}|d_i=1,s_{2i}\big) \tag{62}$$

subject to the constraint

$$\frac{1}{n}\sum_{i=1}^{n}(1-d_i)\max_{p(x_{1i}|d_i=0,s_{1i})} I\big(X_{1i};Y_{1i}|d_i=0,s_{1i}) \\ =\frac{1}{n}\sum_{i=1}^{n}d_i\max_{p(x_{2i}|d_i=1,s_{2i})} I\big(X_{2i};Y_{2i}|d_i=1,s_{2i}), \tag{63}$$

where $n\to\infty$. The maximization of (61) or (62) subject to constraint (63) can be solved by using the Lagrange method as done in [5, Appendix C]. Therefore, we use the proof in [5, Appendix C] by replacing $S(i)$ by

$$S(i)=\max_{p(x_{1i}|d_i=0,s_{1i})} I(X_{1i};Y_{1i}|d_i=0,s_{1i}) \tag{64}$$

and $R(i)$ by

$$R(i)=\max_{p(x_{2i}|d_i=1,s_{2i})} I(X_{2i};Y_{2i}|d_i=1,s_{2i}). \tag{65}$$

Then, the solution of the optimal d_i is obtained as

$$d_i=\begin{cases}1, & \text{if} \quad \max\limits_{p(x_{2i}|d_i=1,s_{2i})} I(X_{2i};Y_{2i}|d_i=1,s_{2i})\geq\rho\times\max\limits_{p(x_{1i}|d_i=0,s_{1i})} I(X_{1i};Y_{1i}|d_i=0,s_{1i})\\ \\ 0, & \text{if} \quad \max\limits_{p(x_{2i}|d_i=1,s_{2i})} I(X_{2i};Y_{2i}|d_i=1,s_{2i})\leq\rho\times\max\limits_{p(x_{1i}|d_i=0,s_{1i})} I(X_{1i};Y_{1i}|d_i=0,s_{1i}),\end{cases} \tag{66}$$

where ρ is a constant found as the solution to constraint (63). d_i in (66) is not unique when

$$\max_{p(x_{2i}|d_i=1,s_{2i})} I(X_{2i};Y_{2i}|d_i=1,s_{2i})=\rho\times\max_{p(x_{1i}|d_i=0,s_{1i})} I(X_{1i};Y_{1i}|d_i=0,s_{1i}) \tag{67}$$

holds[2], and can be either 0 or 1, as long as constraint (63) is satisfied. In order to make d_i unique, we choose to flip a coin each time (67) occurs. For this purpose, we introduce the set of possible outcomes of the coin flip, $\mathcal{C}\in\{0,1\}$, and denote the probabilities of the outcomes

[2]This was not an issue in [5] since for time-continuous fading the probability that (67) holds is negligible.

by $P_{\mathcal{C}} = \Pr\{\mathcal{C} = 1\}$ and $\Pr\{\mathcal{C} = 0\} = 1 - P_{\mathcal{C}}$, respectively. Then, we write (66) as

$$d_i = \begin{cases} 1, \quad \text{if} \quad \max\limits_{p(x_{2i}|d_i=1,s_{2i})} I(X_{2i};Y_{2i}|d_i=1,s_{2i}) > \rho \times \max\limits_{p(x_{1i}|d_i=0,s_{1i})} I(X_{1i};Y_{1i}|d_i=0,s_{1i}) \\ 0, \quad \text{if} \quad \max\limits_{p(x_{2i}|d_i=1,s_{2i})} I(X_{2i};Y_{2i}|d_i=1,s_{2i}) < \rho \times \max\limits_{p(x_{1i}|d_i=0,s_{1i})} I(X_{1i};Y_{1i}|d_i=0,s_{1i}) \\ 1, \quad \text{if} \quad \max\limits_{p(x_{2i}|d_i=1,s_{2i})} I(X_{2i};Y_{2i}|d_i=1,s_{2i}) = \rho \times \max\limits_{p(x_{1i}|d_i=0,s_{1i})} I(X_{1i};Y_{1i}|d_i=0,s_{1i}) \\ \qquad\qquad \text{AND } \mathcal{C} = 1 \\ 0, \quad \text{if} \quad \max\limits_{p(x_{2i}|d_i=1,s_{2i})} I(X_{2i};Y_{2i}|d_i=1,s_{2i}) = \rho \times \max\limits_{p(x_{1i}|d_i=0,s_{1i})} I(X_{1i};Y_{1i}|d_i=0,s_{1i}) \\ \qquad\qquad \text{AND } \mathcal{C} = 0. \end{cases} \tag{68}$$

The constant ρ and the probability $P_{\mathcal{C}}$ are chosen such that constraint (63) is satisfied. Hence, for the optimal d_i given by (68), (60) is equivalent to

$$R \leq \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (1-d_i) \max_{p(x_{1i}|d_i=0,s_{1i})} I(X_{2i};Y_{1i}|d_i=0,s_{1i}) \tag{69a}$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} d_i \max_{p(x_{2i}|d_i=1,s_{2i})} I(X_{2i};Y_{2i}|d_i=1,s_{2i}). \tag{69b}$$

Since the random processes S_1^n and S_2^n are ergodic and stationary, the value of d_i is only dependent on the channel states s_1 and s_2 during the i-th channel use and not on i itself. Hence, d_i for all i for which $S_{1i} = s_1$ and $S_{2i} = s_2$ has the same value and can be written as $d(s_1, s_2)$. As a result, (68) can be written as in (5). Now, given the PMFs of (S_1, S_2) and the probability $P_{\mathcal{C},\text{opt}}$, we can calculate the probability (or the fraction of time) in which states s_1 and s_2 occur and $d(s_1, s_2) = 0$. We denote this probability as $p(d(s_1, s_2) = 0, s_1, s_2)$. Furthermore, we can calculate the probability that states s_1 and s_2 occur and $d(s_1, s_2) = 1$. We denote this probability as $p(d(s_1, s_2) = 1, s_1, s_2)$. Then, the right hand side of (69a) and (69b) can be written as (6) and (7), respectively. This leads to writing (69) as $R \leq C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C$ This concludes the proof of the converse.

Finally, we emphasize that the $(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ for which $C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ occurs is unique. To prove this, assume that for a pair $(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$, $C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ is achieved. Then, let $\mathcal{A}_0$ and $\mathcal{A}_1$ denote sets with pairs of states (s_1, s_2) for which $d(s_1, s_2) = 0$ and $d(s_1, s_2) = 1$, respectively, without using the coin flip. Let $\mathcal{A}_{\mathcal{C}}(\rho_{\text{opt}})$ denote the set of states

for which the coin flip is used. Then each element in $\mathcal{A}_\mathcal{C}(\rho_{\rm opt})$ will produce $d(s_1,s_2)=0$ with probability $1-P_{\mathcal{C},{\rm opt}}$ and $d(s_1,s_2)=1$ with probability $P_{\mathcal{C},{\rm opt}}$. If now ρ is increased, i.e, $\rho>\rho_{\rm opt}$, all of the elements in $\mathcal{A}_\mathcal{C}(\rho_{\rm opt})$ will be transferred to $\mathcal{A}_0$ and for some of the elements in $\mathcal{A}_1$ a coin flip maybe required. Therefore, for the new ρ and any $P_\mathcal{C}$, $C_1(\rho,P_\mathcal{C})>C_1(\rho_{\rm opt},P_{\mathcal{C},{\rm opt}})$ and $C_2(\rho,P_\mathcal{C})<C_2(\rho_{\rm opt},P_{\mathcal{C},{\rm opt}})$ will hold since by increasing ρ we have increased the set $\mathcal{A}_0$ and decreased the set $\mathcal{A}_1$. A similar result is obtained if ρ is decreased, i.e., $\rho<\rho_{\rm opt}$. This concludes the proof of the uniqueness of $(\rho_{\rm opt},P_{\mathcal{C},{\rm opt}})$. ■

B. Achievability of the Capacity

The achievability of the capacity for AWGN channels with slow time-continuous fading was presented in [5]. Here, we outline the same method for achieving the capacity via a buffer, but now for general state-dependent half-duplex channels. However, we are only able to prove achievability when the channel states change slowly such that during each state the channel can be used k times where $k\to\infty$ holds. Furthermore, we show that the capacity can be achieved with a decode-and-forward relay. This is consistent with the results in [1] since a relay channel without a source-destination link is a degraded relay channel, see definition of a degraded relay channel in [1], for which the decode-and-forward operation at the relay was shown to be capacity achieving.

We want to transfer nR bits of information in n channel uses. To this end, the information is sent in N blocks, where $N\to\infty$ and during each block we transmit k symbols, where $k\to\infty$. Obviously, $n=Nk$ has to hold. Furthermore, the relay is equipped with a buffer of unlimited size. For each state $s_1\in\mathcal{S}_1$ of the source-relay channel, assign a rate $R_1(s_1)$ and for each state $s_2\in\mathcal{S}_2$ of the relay-destination channel, assign a rate $R_2(s_2)$. The values of $R_1(s_1)$ and $R_2(s_2)$ will be provided later, cf. (70) and (71), respectively. For each $s_1\in\mathcal{S}_1$ map each combination of $kR_1(s_1)$ bits to a codeword $x_1(s_1)$, and for each $s_2\in\mathcal{S}_2$ map each combination of $kR_2(s_2)$ bits to a codeword $x_2(s_2)$. The codewords $x_1(s_2)$ and $x_2(s_2)$ are such that both are comprised of k symbols and all of the symbols in $x_1(s_1)$ and $x_2(s_2)$ are independent and identically distributed according to $p(x_1|s_1,d(s_1,s_2)=0)$ and $p(x_2|s_2,d(s_1,s_2)=1)$, respectively. Hence, by mapping each combination of $kR_1(s_1)$ and $kR_2(s_2)$ bits to unique codewords, we generate codebooks comprised of $2^{kR_1(s_1)}$ and $2^{kR_2(s_2)}$ codewords, respectively, for each $s_1\in\mathcal{S}_1$ and $s_2\in\mathcal{S}_2$, respectively. Furthermore, before the transmission starts, using the distribution of the random

processes S_1^n and S_2^n, ρ_{opt}, $P_{\mathcal{C},\text{opt}}$, and $d(s_1, s_2)$ can be computed for each pair of $(s_1, s_2) \in \mathcal{S}_1 \times \mathcal{S}_1$ such that $C_1(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) = C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}})$ is achieved, where $C_1(\rho, P_{\mathcal{C}})$ and $C_2(\rho, P_{\mathcal{C}})$ are given by (6) and (7), respectively. Moreover, the pairs of all possible states $(s_1, s_2) \in \mathcal{S}_1 \times \mathcal{S}_2$, can be divided into two groups. One group is those $(s_1, s_2) \in \mathcal{S}_1 \times \mathcal{S}_2$ for which $d(s_1, s_2)$, computed according to (5), does not need a coin flip and therefore, given the respective (s_1, s_2) is uniquely known as 0 or 1. For this group, the relay can generate a lookup table mapping each (s_1, s_2) to a value of $d(s_1, s_2)$ before the start of the transmission. The second group constitutes all pairs of states which do not belong to the first group. These are the states for which a coin flip is used and thereby the value of $d(s_1, s_2)$ is identical to the outcome of the coin flip $\mathcal{C} \in \{0, 1\}$. During the transmission, when a pair of states (s_1, s_2) belonging to the second group occurs, the relay makes a coin flip and sets $d(s_1, s_2) = \mathcal{C}$.

The transmission is carried out in the following way. In the beginning of each block i, the states s_1, (s_1, s_2), and s_2 are acquired by the source, relay, and destination, respectively, where only the relay node can obtain the value of $d(s_1, s_2)$ either by the lookup table or by the coin flip. Then, the relay transmits the value of $d(s_1, s_2)$ using one bit of information to the other nodes. Note that one bit of information is negligible compared to $kR_1(s_1)$ and $kR_2(s_2)$ bits of information transmitted in each block, for $k \to \infty$. Then, if $d(s_1, s_2) = 0$, the source maps $kR_1(s_1)$ bits of information to a codeword $x_1(s_1)$ and transmits the codeword to the relay. The rate $R_1(s_1)$ is given by

$$R_1(s_1) = \max_{p(x_1|d(s_1,s_2)=0,s_1)} I\big(X_1; Y_1|d(s_1, s_2) = 0, s_1\big) - \epsilon, \tag{70}$$

where $\epsilon > 0$ is an arbitrary small number. The relay can decode this codeword since

$$R_1(s_1) < \max_{p(x_1|d(s_1,s_2)=0,s_1)} I\big(X_1; Y_1|d(s_1, s_2) = 0, s_1\big)$$

and store the information bits in its buffer. Let $kQ(i)$ denote the number of bits in the relay's buffer at the end of time slot i. On the other hand, if $d(s_1, s_2) = 1$ the relay extracts $kR_2(s_2)$ bits from its buffer, maps them to a codeword $x_2(s_2)$, and transmits the codeword to the destination. The rate $R_2(s_2)$ is given by

$$R_2(s_2) = \min\left\{Q(i-1)\ \ ,\ \max_{p(x_2|d(s_1,s_2)=1,s_2)} I\big(X_2; Y_2|d(s_1, s_2) = 1, s_2\big)\right\} - \epsilon, \tag{71}$$

where $kQ(i-1)$ is the number of information bits in the relay's buffer at the end of block $i-1$. Note that rate $R_2(s_2)$ is limited by $Q(i-1)$ since the relay cannot transmit more information

than what is in its buffer. The destination can decode this codeword since

$$R_2(s_2) < \max_{p(x_2|d(s_1,s_2)=1,s_2)} I\big(X_2;Y_2|d(s_1,s_2)=1,s_2\big).$$

Using the ergodicity of the processes S_1^n and S_2^n, the average rate of the source during all $i=1,...,N$ blocks is $\sum_{i=1}^N (1-d(s_1,s_2))R_1(s_1)/N$, and for $N\to\infty$ it converges to

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\big(1-d(s_1,s_2)\big)R_1(s_1)\\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\big(1-d(s_1,s_2)\big)\left[\max_{p(x_1|d(s_1,s_2)=0,s_1)} I\big(X_1;Y_1|d(s_1,s_2)=0,s_1\big)-\epsilon\right]\\
&=C_1(\rho_{\text{opt}},P_{\mathcal{C},\text{opt}})-\hat{\epsilon}
\end{aligned}\tag{72}$$

where $C_1(\rho_{\text{opt}},P_{\mathcal{C},\text{opt}})$ is given by (6) and $\hat{\epsilon}=\sum_{i=1}^N(1-d_i)\epsilon/N\le\epsilon$. On the other hand, the average rate of the relay during all $i=1,...,N$ blocks is $\sum_{i=1}^N d(s_1,s_2)R_2(s_2)/N$, and for $N\to\infty$ it converges to

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}d(s_1,s_2)R_2(s_2)\\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}d(s_1,s_2)\left[\min\left\{Q(i-1)\ ,\ \max_{p(x_2|d(s_1,s_2)=1,s_2)} I\big(X_2;Y_2|d(s_1,s_2)=1,s_2\big)\right\}-\epsilon\right].
\end{aligned}\tag{73}$$

Note that (72) is the average arriving rate into the buffer and (73) is the average departing rate out of the buffer. It was proven in [5, Appendix B] that if $d(s_1,s_2)$ is such that

$$\begin{aligned}
&\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\big(1-d(s_1,s_2)\big)\max_{p(x_1|d(s_1,s_2)=0,s_1)} I\big(X_1;Y_1|d(s_1,s_2)=0,s_1\big)\\
&=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}d(s_1,s_2)\max_{p(x_2|d(s_1,s_2)=1,s_2)} I\big(X_2;Y_2|d(s_1,s_2)=1,s_2\big)
\end{aligned}\tag{74}$$

holds, then during all N blocks, the number of blocks, denoted by Δ, in which the buffer does not have enough information stored and thereby

$$\min\left\{Q(i-1)\ ,\ \max_{p(x_2|d(s_1,s_2)=1,s_2)} I\big(X_2;Y_2|d(s_1,s_2)=1,s_2\big)\right\}=Q(i-1)\tag{75}$$

occurs, are negligible compared to the remaining number of blocks $N - \Delta$ in which the buffer does have enough information stored and thereby

$$\min\left\{Q(i-1)\ \ ,\ \ \max_{p(x_2|d(s_1,s_2)=1,s_2)} I\big(X_2;Y_2|d(s_1,s_2)=1,s_2\big)\right\}$$
$$= \max_{p(x_2|d(s_1,s_2)=1,s_2)} I\big(X_2;Y_2|d(s_1,s_2)=1,s_2\big) \tag{76}$$

occurs. As a result, $\lim_{N\to\infty} \Delta/N = 0$ and $\lim_{N\to\infty}(N-\Delta)/N = 1$, and thereby (73) converges to

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} d(s_1,s_2)\left[\max_{p(x_2|d(s_1,s_2)=1,s_2)} I\big(X_2;Y_2|d(s_1,s_2)=1,s_2\big) - \epsilon\right]$$
$$= C_2(\rho_{\text{opt}}, P_{\mathcal{C},\text{opt}}) - \hat{\epsilon}, \tag{77}$$

where $\hat{\epsilon} \le \epsilon$. This concludes the proof of the achievability of the capacity.

V. Conclusion

We have derived the capacity of the state-dependent half-duplex relay channel without source-destination link. We have proven the converse, the achievability, and applied the capacity formula to three specific channels. Thereby, we have shown that the a half-duplex relay channel offers a degree of freedom which has been previously overlooked. This is the freedom of the half-duplex relay to choose either to transmit or receive depending on the quality of its respective receiving and transmitting channels.

References

[1] T. Cover and A. El Gamal, "Capacity Theorems for the Relay Channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572–584, Sep. 1979.

[2] M. Khojastepour, A. Sabharwal, B. Aazhang *et al.*, "On the Capacity of Cheap Relay Networks," in *Conference on Information Sciences and Systems*, 2003.

[3] A. Host-Madsen and J. Zhang, "Capacity Bounds and Power Allocation for Wireless Relay Channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2020 –2040, Jun. 2005.

[4] M. Khojastepour, "Distributed Cooperative Communications in Wireless Networks," Ph.D. dissertation, Rice University, 2004.

[5] N. Zlatanov, R. Schober, and P. Popovski, "Buffer-Aided Relaying with Adaptive Link Selection," *Accepted for publication in IEEE J. Select. Areas Commun.*, 2012. [Online]. Available: http://arxiv.org/abs/1202.5349

[6] N. Zlatanov and R. Schober, "Buffer-Aided Relaying with Adaptive Link Selection - Fixed and Mixed Rate Transmission," *Accepted for publication in IEEE Trans. Inform. Theory*, 2012. [Online]. Available: http://arxiv.org/abs/1207.4763

[7] R. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[8] A. El Gamal and Y. Kim, *Network Information Theory*. Cambridge University Press, 2011.